APPENDIX B
Survey Methodology

Sample Design

A Stratified Two-stage Sampling was adopted for the survey. All provinces (or changwat) were considered to be constituted strata, totally 76 strata. Each stratum was divided into two parts according to the type of local administration, namely municipal areas and non-municipal areas. The primary sampling units were blocks for municipal areas and were villages for non-municipal areas. The secondary sampling units were private households.

Selection of Sampling Unit

As this survey was an unscheduled project, the sample size was limited only 16,500 households and data collection was only 6 months (March-May and August-October 2001). To avoid a work load or a delay of the regular plan of NSO, all sampled blocks / villages used in March - May and August - October 2000 were primary sample employed in the same period of field operation in 2001 in order to obtaining information which could reflect the economic crisis on households explicitly. Therefore, there were altogether 1,690 sampled blocks / villages distributed according to regions and administrative areas as follows;

Region	Total	Municipal Areas	Non-Municipal Areas
Greater Bangkok Metropolitan Area*	172	126	46
Central (Excluding Greater Bangkok Metropolitan Area)	416	208	208
North	374	188	186
Northeast	442	228	214
South	286	144	142
Total	**1,690**	**894**	**796**

* Greater Bangkok Metropolitan Area, i.e., Bangkok Metropolis, Nonthaburi province, Pathum Thani province and Samut Prakan province.

For the second stage of selecting private household, the 1999 – 2000 list of basic household characteristics of sample blocks / villages was employed. A new set of sampled households was selected systematically. The sample size was as follows;

    (1)   15 households from each of sample blocks in municipal areas,

          9 households from each of sample blocks in sanitary districts which were reclassified as municipal areas,

    (2)   7 households from each of the sample villages.

Before selecting the private sample households, the set of households was rearranged by size of household (identified by number of household members) and type of economic household. The total number of private sampled households which were enumerated during March – May and August – October 2001 by region and administrative areas was as follows;

Region / Stratum	Total	Municipal Areas	Non-Municipal Areas
Greater Bangkok Metropolitan Area*	2,104	1,782	322
Central (Excluding Greater Bangkok Metropolitan Area)	3,916	2,460	1,456
North	3,522	2,220	1,302
Northeast	4,162	2,664	1,498
South	2,710	1,716	994
Total	**16,414**	**10,842**	**5,572**

*   Greater Bangkok Metropolitan Area, i.e., Bangkok Metropolis, Nonthaburi province, Pathum Thani province and Samut Prakan province.

Method of Estimation

The survey results were presented at regional levels and administrative areas. The method of estimated characteristics were as follows :

Let

$j = $ the serial number of block/village $\quad j = 1, 2, 3,...m_{hki}$

$i = $ Type of local administration $\quad i = \begin{cases} 1 & \text{for municipal areas} \\ 2 & \text{for non-municipal areas} \end{cases}$

$k = $ province $\quad\quad k = 1, 2, 3,...A_h$

$h = $ region $\quad\quad\quad h = \begin{cases} 1 & \text{for Greater Bangkok} \\ & \text{Metropolis Area} \\ 2 & \text{for Central region (excluding} \\ & \text{Greater Bangkok Metropolis} \\ & \text{Area)} \\ 3 & \text{for Northern region} \\ 4 & \text{for Northeastern region} \\ 5 & \text{for Southern region} \end{cases}$

1) Estimation of Total

(1) Estimate of the Total Number of Characteristic Y of Household

1.1 Adjusted estimate of the total number of characteristic Y of household for the ith area, kth province, hth region was based on the formula :

$$Y''_{hki} = \frac{Y'_{hki}}{H'_{hki}} H''_{hki} = r_{hki} H''_{hki} \qquad \text{…................. (1)}$$

where $\quad Y'_{hki}$ is the ordinary estimate of the total number of characteristic Y of household in the ith area, kth province, hth region.

H'_{hki} is the ordinary estimate of the total number of households in the ith area, kth province, hth region.

H''_{hki} is the estimate, based on the population projection, of the total number of households [1] in the ith area, kth province, hth region.

---

[1] Population Projections for Thailand 1990 - 2020, National Economic and Social Development Board

r_{hki} is the ratio of the estimate of the total number of characteristic Y to the estimate of the total number of households in the ith area, k^{th} province, h^{th} region.

The formula of the estimate from a stratified two - stage sampling was as follows :

$i)$
$$Y'_{hki} = \frac{1}{m_{hki}} \sum_{j=1}^{m_{hki}} \frac{1}{P_{hkij}} \frac{N_{hkij}}{n_{hkij}} y_{hkij} \qquad \text{.................... (2)}$$

where y_{hkij} is the characteristic Y of every sample household in the jth sample block / village, i^{th} area, k^{th} province, h^{th} region.

N_{hkij} is the total number of listing households in the jth sample block / village, i^{th} area, k^{th} province, h^{th} region.

n_{hkij} is the total number of sample households in the jth sample block/ village, i^{th} area, k^{th} province, h^{th} region.

P_{hkij} is the probability of selection of the jth sample block / village, i^{th} area, k^{th} province, h^{th} region.

m_{hki} is the total number of sample blocks / villages in the ith area, k^{th} province, h^{th} region.

$ii)$
$$H'_{hki} = \frac{1}{m_{hki}} \sum_{j=1}^{m_{hki}} \frac{1}{P_{hkij}} \frac{N_{hkij}}{n_{hkij}} n'_{hkij} \qquad \text{.................... (3)}$$

where n'_{hkij} is the total number of the interviewed households in the jth sample block / village, i^{th} area, k^{th} province, h^{th} region.

1.2 Adjusted estimate of the total number of characteristic Y of household for the ith area, hth region was based on the formula :

$$Y''_{hi} = \sum_{k=1}^{A_h} Y''_{hki} \qquad \text{…................. (4)}$$

where A_h is the total number of provinces in the hth region and

$$\sum_{h=1}^{5} A_h = 76$$

1.3 Adjusted estimate of the total number of characteristic Y of household for the hth region was based on the formula :

$$Y''_h = \sum_{i=1}^{3} Y''_{hi} \qquad \text{…................. (5)}$$

1.4 Adjusted estimate of the total number of characteristic Y of household for the ith area of the whole kingdom was based on the formula :

$$Y''_i = \sum_{h=1}^{5} Y''_{hi} \qquad \text{…................. (6)}$$

1.5 Adjusted estimate of the total number of characteristic Y of household for the whole kingdom was based on the formula :

$$Y'' = \sum_{h=1}^{5} Y''_h = \sum_{i=1}^{3} Y''_i \qquad \text{.................... (7)}$$

(2) Estimate of Variance of the Total Number of Characteristic Y of Household

2.1 The estimate variance of Y''_{hki} was

$$\hat{V}(Y''_{hki}) = \left[\frac{H''_{hki}}{H'_{hki}} \right]^2 \frac{1}{m_{hki}(m_{hki} - 1)} \sum_{j=1}^{m_{hki}} z^2_{hkij} \qquad \text{…................. (8)}$$

where $z_{hkij} = Y'_{hkij} - r_{hki}H'_{hkij}$

$$Y'_{hkij} = \frac{1}{P_{hkij}}\frac{N_{hkij}}{n_{hkij}}Y_{hkij}$$

$$H'_{hkij} = \frac{1}{P_{hkij}}\frac{N_{hkij}}{n_{hkij}}n'_{hkij}$$

2.2 The estimate variance of Y''_{hi} was

$$\hat{V}(Y''_{hi}) = \sum_{k=1}^{A_h} \hat{V}(Y''_{hki})$$

…................. (9)

2.3 The estimate variance of Y''_h was

$$\hat{V}(Y''_h) = \sum_{i=1}^{3} \hat{V}(Y''_{hi})$$

.... (10)

2.4 The estimate variance of Y''_i was

$$\hat{V}(Y''_i) = \sum_{h=1}^{5} \hat{V}(Y''_{hi})$$

…................. (11)

2.5 The estimate variance of Y'' was

$$\hat{V}(Y'') = \sum_{h=1}^{5} \hat{V}(Y''_h) = \sum_{i=1}^{3} \hat{V}(Y''_i)$$

…................. (12)

(3) Coefficient of Variation (CV) of the Total Number of Characteristic Y of Household

3.1 The formula of CV of Y''_{hi} was

$$CV(Y''_{hi}) = \frac{\sqrt{\hat{V}(Y''_{hi})}}{Y''_{hi}} \times 100 \text{ \%} \qquad \text{…................. (13)}$$

3.2 The formula of CV of Y''_h was

$$CV(Y''_h) = \frac{\sqrt{\hat{V}(Y''_h)}}{Y''_h} \times 100 \text{ \%} \qquad \text{…................. (14)}$$

3.3 The formula of CV of Y''_i was

$$CV(Y''_i) = \frac{\sqrt{\hat{V}(Y''_i)}}{Y''_i} \times 100 \text{ \%} \qquad \text{…................. (15)}$$

3.4 The formula of CV of Y'' was

$$CV(Y'') = \frac{\sqrt{\hat{V}(Y'')}}{Y''} \times 100 \text{ \%} \qquad \text{…................. (16)}$$

2) Estimation of Average

(1) Estimation of the Average of Characteristic Y per Household

1.1 The estimate of the average of characteristic Y per household for the ith area,

hth region was based on the formula :

$$\overline{Y}_{hi} = \frac{Y''_{hi}}{H''_{hi}} = r_{hi} \qquad \text{…….............. (17)}$$

where $\qquad H''_{hi} = \sum_{k=1}^{A_h} H''_{hki}$

1.2 The estimate of the average of characteristic Y per household for the hth

region was based on the formula :

$$\overline{Y}_h = \frac{Y''_h}{H''_h} = r_h \qquad \text{…................ (18)}$$

where $\quad H''_h = \sum_{i=1}^{3} H''_{hi}$

1.3 The estimate of the average of characteristic Y per household for the ith area

of the whole kingdom was based on the formula :

$$\overline{Y}_i = \frac{Y''_i}{H''_i} = r_i \qquad \text{…................ (19)}$$

where $\quad H''_i = \sum_{h=1}^{5} H''_{hi}$

1.4 The estimate of the average of characteristic Y per household for the whole

kingdom was based on the formula :

$$\overline{Y} = \frac{Y''}{H''} = r \qquad \text{…................ (20)}$$

where $\quad H'' = \sum_{h=1}^{5} H''_h = \sum_{i=1}^{3} H''_i$

(2) Estimate of Variance of the Average of Characteristic Y per Household

2.2 The estimate variance of $\overline{Y}_{hi}$ was

$$\hat{V}(\overline{Y}_{hi}) = \left[\frac{1}{H'_{hi}}\right]^2 \sum_{k=1}^{A_h} \frac{1}{m_{hki}(m_{hki}-1)} \left[\sum_{j=1}^{m_{hki}} z''^2_{hkij} - m_{hki} z'^2_{hki}\right] \qquad \text{......(21)}$$

where $\quad z''_{hkij} = Y'_{hkij} - r_{hi} H'_{hkij}$

$\qquad\qquad z'_{hki} = Y'_{hki} - r_{hi} H'_{hki}$

2.2 The estimate variance of $\overline{Y}_h$ was

$$\hat{V}(\overline{Y}_h) = \left[\frac{1}{H'_h}\right]^2 \sum_{i=1}^{3} \sum_{k=1}^{A_h} \frac{1}{m_{hki}(m_{hki}-1)} \left[\sum_{j=1}^{m_{hki}} z'''^2_{hkij} - m_{hki} z''^2_{hki}\right] \; .. \, (22)$$

where $\quad z'''_{hkij} = Y'_{hkij} - r_h H'_{hkij}$

$\qquad\qquad z''_{hki} = Y'_{hki} - r_h H'_{hki}$

2.3 The estimate variance of $\overline{Y}_i$ was

$$\hat{V}(\overline{Y}_i) = \left[\frac{1}{H'_i}\right]^2 \sum_{h=1}^{5} \sum_{k=1}^{A_h} \frac{1}{m_{hki}(m_{hki}-1)} \left[\sum_{j=1}^{m_{hki}} z''''^2_{hkij} - m_{hki} z'''^2_{hki}\right] \; .. \, (23)$$

where $\quad z''''_{hkij} = Y'_{hkij} - r_i H'_{hkij}$

$\qquad\qquad z'''_{hki} = Y'_{hki} - r_i H'_{hki}$

2.4 The estimate variance of $\overline{Y}$ was

$$\hat{V}(\overline{Y}) = \left[\frac{1}{H'}\right]^2 \sum_{h=1}^{5} \sum_{i=1}^{3} \sum_{k=1}^{A_h} \frac{1}{m_{hki}(m_{hki}-1)} \left[\sum_{j=1}^{m_{hki}} z'''''^2_{hkij} - m_{hki} z''''^2_{hki}\right] \; ... \, (24)$$

where $\quad z'''''_{hkij} = Y'_{hkij} - r H'_{hkij}$

$\qquad\qquad z''''_{hki} = Y'_{hki} - r H'_{hki}$

(3) Coefficient of Variation (CV) of the Average of Characteristic Y per Household

3.1 The formula of CV of $\bar{Y}_{hi}$ was

$$CV(\bar{Y}_{hi}) = \frac{\sqrt{\hat{V}(\bar{Y}_{hi})}}{\bar{Y}_{hi}} \times 100 \ \% \qquad\qquad \text{.................. (25)}$$

3.2 The formula of CV of $\bar{Y}_{h}$ was

$$CV(\bar{Y}_{h}) = \frac{\sqrt{\hat{V}(\bar{Y}_{h})}}{\bar{Y}_{h}} \times 100 \ \% \qquad\qquad \text{................... (26)}$$

3.3 The formula of CV of $\bar{Y}_{i}$ was

$$CV(\bar{Y}_{i}) = \frac{\sqrt{\hat{V}(\bar{Y}_{i})}}{\bar{Y}_{i}} \times 100 \ \% \qquad\qquad \text{................... (27)}$$

3.4 The formula of CV of $\bar{Y}$ was

$$CV(\bar{Y}) = \frac{\sqrt{\hat{V}(\bar{Y})}}{\bar{Y}} \times 100 \ \% \qquad\qquad \text{................... (28)}$$

Data Collection and Data Processing

    1)   Data Collection

        (1) Data Collection Method   The survey data was collected by an interviewing method. The interviewers from the Central Office and the NSO provincial branch offices were sent out to interview the heads of the sample households or other household members. The information obtained was recorded on the questionnaire forms. Two types of questionnaire forms were used, the SES 2 for household composition, demographic and economic characteristics of household members, income and housing facilities, and the SES 3 for household expenditures and 7‐day food consumption.

        (2) Collection Period   The interviews were conducted on the time period of six months (March – May and August – October 2001), so all sample households were divided into six equally representative sub-samples. Each sub-household group was interviewed for the period of one-month.

(3) Data Reference Period

    (3.1)   Data for goods and services expenditure was obtained for the month preceding the month in which the interview was conducted. For example, in an interview result of March 2001, "the previous month" referred to 1– 28 February 2001.

    (3.2)   Data for expenditure on items not frequently purchased was obtained for the twelve months before the month in which the interview was conducted. For instance, in an interview result of March 2001, "during the past twelve months" referred to the period between March 2000 to February 2001, and to derive the average monthly expenditure, those amount were divided by 12.

    (3.3)   Household income was the data of the past 12 months and averaged out in household monthly income.

(4) Quality Control   To ensure the highest quality of data, each completed questionnaire was subjected to thorough field editing, followed by a follow-up interview if the information was found to be incomplete or internally inconsistent. Moreover, a household account balance sheet was prepared for each completed interview.   This balance compared total money "disbursements" with total money "receipts" for the preceding month. If the account was more than 15 percent out of balance, the interviewer was expected to revisit the household to reconcile the difference.

During the survey period, the supervisors were assigned to assist interviewers in solving the arising problems. Moreover, members of the Central Office staff conducted periodic visits to the field to review questionable reports and clarify data collection procedures.

2) Data Processing

(1) Editing and coding   In the editing and coding process, all questionnaires were examined for completeness and consistency by the supervisors. Descriptive information was coded numerically for computer processing.   All annual expenditure and income values were converted to a one month base by dividing annual values by 12. For 7‑day food consumption, values were multiplied by 4.3, which was the average number of weeks per month (52 weeks / 12 months = 4.3 weeks/month). Income from farm or non‑farm enterprises was calculated on the basis of total annual value of production, less operating expenses.   From this estimate the value of products held or withdrawn for household consumption was subtracted to arrive at an estimate of money income.

At the Central Office, questionnaires were again reviewed for completeness and internal consistency before they were sent to the Data Processing Operations Division.

(2) Tabulation   All information to be tabulated was keyed to tape for computer processing. Before tabulations were prepared, all raw data was edited by computer for final review and correction to ensure that all the data was 'clean,' ready for further tabulation.

The data presented in this report is in this form of averages and percentage distributions. Average monthly expenditure and income,and average amount of debt per household for the survey year are based on all households falling in a specified class regardless of whether the household incurred an expense for a given item or received income from a given source, or having or not having debt. An individual household may have spent much more or much less than the average, or nothing at all. Even among households having similar characteristics, expenditure and income may differ substantially. It is expected however that variations in the average expenditure and income patterns will reflect significant differences between households living in different circumstances and at different levels of living.

Limitations of the Data

Sample surveys are subject to various types of errors. Sampling errors occur because observations are not taken from the entire population. Non-sampling errors can be attributable to many sources, such as inability to obtain information from all households selected in the sample, inability or unwillingness of respondent households to provide correct information, errors made in recording data, mistakes made in coding and estimating for missing data. Every effort was made to minimize errors of all types.

Generally, income and expenditure data are under-reported by the sample households. Income data is found to be under-reported more than expenditures. The degree of under-reporting varies according to different sources of income. For example, wage and salary earnings are probably much more accurately reported than profits from business. Households are likely to forget to report some expenditures made during the reference period of the survey and household members may be reluctant to report certain types of expenditures, for example, on alcoholic drinks, cigarettes and entertainment activities.

Income-in-kind was imputed by respondent. Self-produced commodities produced individually and consumed by the household, and goods received free were valued at current local retail prices. The rental value of an owner-occupied home was the value that would have been received per month had the house been rented out to another household.

In the 2001 Special Survey, the collection of data was conducted for six months (March – May and August – October 2001). It was different from the regular one which was the twelve months basis and the sampling units of the survey was about 16,500 households or a half of the total sampled households of the regular one. The data obtained from this survey may be distorted from the regular one due to some types of household expenditures being varied from season to season. Comparison of this set of data to the previous ones should be considered cautiously.